

EXECUTIVE SUMMARY

I. Background and Study Objectives

The Centers for Medicare and Medicaid Services (CMS), the Federal agency that administers the Medicare program, contracts with a national network of 53 Quality Improvement Organizations (QIOs)—one in each state, the District of Columbia, Puerto Rico, and the Virgin Islands. QIOs seek to 1) improve the quality of care that Medicare beneficiaries receive by collaborating with providers to help them meet evidence-based standards of care, 2) protect beneficiaries by responding to and investigating claims and evidence of substandard care, and 3) protect the Medicare Trust Funds by reviewing claims patterns and suspicious cases for the inappropriate use of services or incorrect billing codes. Over the course of a 3-year contract with CMS, QIOs engage providers in quality improvement projects and offer technical assistance across four major health care settings – hospitals, home health agencies, nursing homes, and physician offices. For the current 3-year contract period CMS has dedicated \$1.265 billion to the program.

Recent press coverage and inquiries made by Congress have raised questions regarding the QIO program’s effectiveness and whether substantial reforms should be made to the program. As part of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA) of 2003, the Congress requested that the Institute of Medicine (IOM) conduct an evaluation of the QIO program. The IOM released their report “Medicare’s Quality Improvement Organizations: Maximizing Potential” in March 2006. Among the IOM’s conclusions was that:

“Given the lack of consistent and conclusive evidence in scientific literature and the lack of strong findings from the committee’s analyses, it is not possible to determine definitively the extent of the impact of the QIOs and the national QIO infrastructure on the quality of health care received by beneficiaries. Many confounding factors make it difficult to attribute the results obtained thus far [to QIOs].” (IOM, 2006)

I.A Study Objectives

In 2005 ASPE contracted with NORC at the University of Chicago (NORC) to develop several options for evaluating the effectiveness of the QIO program. NORC’s objectives for this study were three-fold:

- 1) Conduct an environmental scan to identify and create an inventory of QIO-specific technical assistance activities, interventions, and strategies used to meet performance targets identified in the 7th and 8th SOW and enter this data into a database of QIO activities;
- 2) Conduct site visits to QIOs to gather more detailed information about their day-to-day operations and quality improvement strategies;
- 3) Identify alternative designs for evaluating the QIO program or studies to enhance our

understanding of selected components of the program, to be vetted by members of a Technical Expert Panel (TEP).

I.B History and Structure of the QIO Program

The origins of the QIO program date back more than thirty years, beginning in 1971 with the creation of Experimental Medical Care Review Organizations (EMCROs), in 1972 with the creation of Professional Standards Review Organizations (PSROs), and then in 1982 with the creation of the Utilization and Quality Control Peer Review Organization (PRO) Program. These earlier programs focused on utilization review, cost-containment, and adherence to local practice patterns by “inspecting and detecting” to identify egregious cases in delivery of care and, if necessary, sanctioning providers for substandard care. As a result, providers perceived them more as adversarial and regulatory in nature, as opposed to potential partners in quality improvement.

In response to a 1990 review by the Institute of Medicine (1990), which concluded that a collaborative approach to quality improvement would be more effective in improving providers’ performance, the Health Care Financing Administration (HCFA) (now CMS) launched the Health Care Quality Improvement Initiative (HCQII) in 1992 to analyze patterns of care and identify areas for improvement. Under the HCQII, PROs were encouraged to collaborate with hospitals as partners in developing and implementing hospital quality improvement initiatives instead of focusing on identifying individual “bad apples” within the provider community. These changes implemented by HCFA represented a dramatic shift in vision for the QIO program. Subsequently, Congress officially renamed the PRO program in 2001 to the “Quality Improvement Organization Program.”

To date, eight rounds of contracting have occurred since the shift to a 3-year contract cycle took place in 1984, hence bringing us in 2005 to the 8th Scope of Work (SOW). Under the SOW QIOs are required to engage in four major sets of tasks. Tasks 1 through 3 are referred to in this report as the “core contract” since all QIOs are required to perform these activities. Task 4 refers to “non-core” activities. These are “Special Studies,” which selected QIOs may be contracted to perform.

Under Task 1 of the 8th SOW core contract, QIOs are responsible for providing technical assistance to providers across four major health care settings – nursing homes, home health agencies, hospitals, and physician offices – in order to improve providers’ performance across multiple clinical outcomes and processes of care measures. Furthermore, CMS requires that QIOs divide their technical assistance activities between two different groups of providers. First, QIOs must offer technical assistance to all providers in a state who request assistance on issues of quality improvement as identified in the SOW. The second group of providers includes an “identified participant group,” or an IPG. Providers in an IPG are selected by QIOs and, subsequently, volunteer to receive intensive and ongoing technical assistance and participate in a number of projects to meet specified performance improvement targets. Thus, Task 1 is comprised of QIOs’ activities with IPG and non-IPG providers. Under Task 3, QIOs review beneficiary complaints for quality of care concerns and, as part of the Hospital Payment Monitoring Program (HPMP), they also review the accuracy of DRG codes, medical necessity, and the appropriateness of care to address issues of inappropriate utilization or billing patterns.

Task 4 of the SOW is comprised of the Special Studies Program. The Special Studies Program includes two different types of special studies—Quality Improvement Organization Support Centers

(QIOSCs) and all other special studies. CMS awards QIOs funds to conduct special studies in addition to their core contract activities. Special studies are designed to gather information for identifying best practices; examining or testing performance measures, tools or technical assistance approaches; and, in general, addressing issues of specific interest or relevance to CMS and the QIO program. Quality Improvement Organization Support Centers are QIOs who receive funds to offer technical assistance or support to other QIOs by providing them with the tools, training, information on best practices, and other resources that they need to work effectively with providers to meet quality improvement objectives. As of the 8th SOW, a total of 15 QIOSC contracts have been awarded.

I.C Review of the Literature on QIO Program Effectiveness

For years, researchers have attempted to evaluate the effectiveness of the QIO program using both qualitative and quantitative analytical techniques and with national-, organizational-, and health care setting-level data, but, for the most part, these studies have proven inconclusive. Even the most recent studies are plagued by the same methodological obstacles that earlier studies failed to overcome – questionable data, selection bias, spurious attribution due to numerous confounding factors (e.g. secular trends, differences in provider motivation, non-QIO quality improvement initiatives), lack of generalizability, and the inability to isolate and define experimental and control groups.

The body of literature on the QIO program brings to policymakers' attention the importance of quality improvement in Medicare and, in part, suggests that QIOs play a role in promoting quality of care. However, the evidence is inconclusive as to what extent, if any, of the demonstrated quality improvement can be attributed to the QIO program, overall. This conclusion stems from two major observations in the literature:

- The review of this literature did not yield a conclusive answer as to whether or not the QIO program or specific QIO-led interventions resulted in higher quality, lower quality or no change in any given provider setting. While several QIO interventions or collaboratives suggest that QIO-directed quality improvement activities have been effective at improving selected process and outcome measures, the statistical significance of the findings varied. As an editorial in a 2005 issue of *JAMA* pointed out that among 33 recent studies of the QIO program, 16 yielded “ambiguous results,” eight reported no or negative effects, and nine reported positive effects.
- Most studies evaluating the effectiveness of the QIO program are fraught with methodological limitations—such as selection bias, confounding, and attribution—that are inherent in the study designs. Such problems are threats to the internal and external validity of the studies and may bias study findings. In the future, new and methodologically rigorous studies will be necessary to offer more meaningful conclusions about the effectiveness of the QIO program.

II. Major Findings from QIO Inventory, Site Visits and TEP Meeting

II.A Development of QIO Inventory

In order to obtain an inventory of QIO activities for the 7th and 8th SOWs, NORC conducted a comprehensive environmental scan. As part of this scan we gathered a standardized set of descriptive information about each of the 53 QIOs; data consisted of basic identifying information such as address and the name of the Chief Executive Officer. Other data consisted of information on the organizational structure, profit status, board membership and composition. To the extent available, we gathered activity-level information on each of the QIOs and information related to the organization's day-to-day operations and activities, such as ongoing quality improvement projects and initiatives; related publications; trainings, workshops, and other services offered to providers; collaborations with other organizations; and beneficiary outreach activities. Information gathered from the environmental scan was used to populate a database or inventory of QIO activities, and to develop QIO-specific site visit interview protocols. Finally, data from the scan assisted staff in the development of evaluation designs.

For the overwhelming majority of tasks, large gaps exist in the data. The scope of findings reflected the paucity of activity- or intervention-specific information available in public resources, particularly activities related to the 7th SOW. In several cases, no substantive information on any specific project could be found for a given QIO and subtask. The quality and depth of information did, nonetheless, vary greatly from QIO to QIO. Even for a single QIO, the information available often varied from setting to setting. Efforts to locate details on projects that were identified by name often proved futile and while most QIOs stated that they currently or have previously participated in national or local quality improvement initiatives, specific details as to the QIOs' scope or role in the initiative were generally unavailable.

II.B Site Visits to QIOs

To gain on-the-ground insight into individual QIOs' daily operations, NORC conducted site visits to nine QIO contractors, representing 12 states and the District of Columbia. In consultation with ASPE and CMS staff, site visit QIOs were chosen on the basis of the size of the state they served, location, whether they held single or multiple QIO contracts or QIOSC contracts, and profit status. QIO staff were queried about organizational structure and governance, their strategies for completing tasks under and beyond the core contract (such as special study and/or QIOSC activities), and their experiences with CMS management of the program, including the contracting and evaluation process. A brief overview of the site visit results is presented below.

Identified participant group selection: Most QIOs report “cherry-picking” in order to meet CMS's performance targets, that is, QIOs choose providers as identified participants who are most likely to garner QIOs a passing score on CMS's evaluation. Moreover, QIOs indicated that they tend to avoid working with both poor performers and high performers – the former because they may lack the resources or the motivation to meet the SOW's quality improvement benchmarks and the latter due to a possible “ceiling effect” that may limit the degree of potential performance improvement.

Technical assistance offered to providers: QIO perceptions of which forms of technical assistance are most effective differed—some preferred collaborative models or group training, while others preferred a consultative approach incorporating one-on-one assistance. QIOs reported that the technical assistance strategies they employ depend, in part, on budgetary constraints, geographic distribution of providers, the presence of field offices, and the type of provider and subtask. Additionally, QIOs reported that increasing micromanagement on the part of CMS and CMS’ data lags have restricted both their ability to innovate in order to better respond to the unique needs of the communities they serve and to conduct real-time tracking of the impact of specific interventions.

Case review and beneficiary protection: All QIOs reported that they receive relatively few beneficiary complaints and, furthermore, they indicated that most complaints received were not true quality of care issues, rather, complaints tended to deal with service problems, such as long wait times, “rude staff,” and other communication problems. Despite this, all QIOs disagreed with the IOM’s recommendation that case review activities be removed from QIOs’ responsibilities.

II.C Proceedings from Technical Expert Panel Meeting

NORC identified and recruited eight experts to respond to and offer feedback and guidance on the draft evaluation design options. The TEP was convened to ensure that the evaluation designs NORC proposed were as rigorous and appropriate as feasible considering the scope of the project, the availability (or lack thereof) of data, and the constraints facing the government and an eventual evaluator of the QIO program. The TEP provided several major recommendations, including:

- Evaluations of the QIO program should be prospective. That is, all necessary data collection vehicles should be in place at the start of the 9th SOW in order to support ongoing evaluation activities throughout the SOW period of performance. Moreover, a prospective evaluation may enable the use of more rigorous methodological techniques, such as randomized case control designs.
- Options for evaluating the program, as a whole, are limited due to a number of methodological barriers, thus, multiple smaller-scale studies may be more feasible, such as well-designed case control studies or randomized control trials to examine the effectiveness of different technical assistance interventions. These types of studies could potentially minimize attribution issues and yield results that are more actionable.
- Several members suggested that instead of the historic snapshot approach to the QIO program evaluation, a shift in paradigm to continuous quality improvement would be more informative and may better enable organizations to shift courses to make necessary programmatic changes.

III. Evaluation Designs and Considerations

This section describes general approaches for evaluating both the core QIO program and supplementary components of the program, including special studies and QIOSC contracts,

and non-evaluative studies that could be used to gather information or develop tools to enhance future evaluations of the QIO program as well as to gain a more refined understanding of the program's role in quality improvement. The proposed evaluation options build on prior evaluations that have been conducted, but uses econometric and statistical approaches to address several of the methodological limitations affecting these studies. We also build upon findings from our QIO inventory and site visits to QIOs. A major resource in shaping our recommendations was the 2006 report "Medicare's Quality Improvement Organization Program: Maximizing Potential," issued by the IOM Committee on Redesigning Health Insurance, Performance Measures, Payment, and Performance Improvement Programs. Finally, the evaluation options described were informed and shaped by the input of an eight-member Technical Expert Panel (TEP).

III.A Designs for Evaluating the Core QIO Program

We begin this discussion by describing a design option that is based on a national, provider-level analysis which incorporates a case-control panel design to assess differences in IPG and non-IPG providers' performance. Limitations to this approach are described in the body of the report.

Long-term evaluation goal and approach: In situations where a randomized control trial cannot be used, a two-stage econometric model may be used to estimate program effects. Thus, we propose using econometric modeling to examine differences in IPG and non-IPG provider performance on clinical quality and process of care measures. It is hypothesized that for each health care setting under Task 1, performance on quality measures (e.g., restraint use in nursing homes, on-time prophylactic antibiotic administration in hospitals, etc.) is related directly to provider engagement with the QIO. This hypothesis, however, is flawed due to the presence of selection bias, that is, there may be inherent differences between providers who were selected (or volunteered) to participate in an IPG and providers who were not selected (or did not volunteer) to participate. Due to non-random selection, and the likelihood that IPG providers are selected to participate because they are the most likely to improve (or they volunteer to participate because they are the most motivated to improve), estimates of a QIO's impact on performance likely will be biased.

A two-stage econometric modeling approach can be used to account for factors that may influence a providers' likelihood of working with a QIO, thereby helping to address the two methodological barriers that have hindered previous QIO program evaluations – selection bias and confounding, or attribution. The first equation models the selection mechanism by estimating the probability that a provider of a particular type (e.g., nursing home, home health agency etc.) participates or is selected to participate in a QIO's IPG. The second equation addresses selection bias by estimating provider performance as a function of the likelihood of selection into an IPG as well as other variables that include provider, environmental, and QIO characteristics.

Primary and secondary data collection activities: Primary and secondary data collection will be required to model the dependent and independent variables that comprise the relationships described above. The major dependent variables are provider participation in an IPG and provider performance on subtask quality measures.

- Provider participation in an IPG. Due to regulations that limit access to data on which providers are IPG members, evaluators must currently work directly through individual QIOs or QIOSCs

to gather de-identified data on IPG providers, or through CMS to obtain access to the PARTner System, which also stores this type of data electronically.

- Provider performance on subtask quality measures. These data are collected as a standard part of the QIO program and should continue to be available through CMS or the QIOSCs. In fact, for many subtasks, the performance measures by which QIO performance is evaluated are the same measures reported publicly in the hospital, home health, and nursing home COMPARE databases or obtained from the Nursing Home Minimum Data Set (MDS) or the Home Health Outcome and Assessment Information Set (OASIS).
- The major independent variables used in this model are provider, environmental, and QIO characteristics. Year or time period also is included in this model because, as suggested by a member of the TEP, an effort should be made to examine continuous improvements in quality. As such, it is recommended that performance be measured on at least an annual basis.
- Provider characteristics. The probability of selection (the first equation in the model) or participation in an IPG could be driven by a number of provider-level characteristics. CMS administrative databases (Providers of Services file, the Medicare Cost Reports, the Standard Analytical Files, and the Provider Enrollment, Chain and Ownership System data) may be used to extract information on provider profit status, membership in a system, rural/urban location, and staffing. Private sector databases, such as the American Hospital Association Annual Survey, may supplement information that is not available in CMS administrative databases. Information on providers' level of motivation and willingness to work with QIOs on quality improvement issues, the extent to which the provider has the internal infrastructure to support quality improvement efforts, and utilization of non-QIO quality improvement resources is not readily available and must be obtained through primary data collection. A potential primary data collection instrument is the CMS "Survey of Provider Satisfaction with Quality Improvement Organizations."
- Environmental characteristics. Certain environmental characteristics may impact providers' willingness to work with QIOs, such as whether providers are required by managed care organizations to participate in selected quality improvement initiatives or the level of market competitiveness. Resources to characterize environmental features that may drive participation in an IPG and other quality improvement activities are available from public and private sources, such as the Bureau of Health Profession's Area Resource File, the Medicare Denominator File (for use in estimating managed care penetration in the elderly population), and the Kaiser Family Foundation State Health Facts database.
- QIO characteristics. It is probable that quality improvement is driven not solely by whether a provider is an IPG member, but also by the types, intensity, and frequency of technical assistance that QIOs offer to providers. The concepts of "technical assistance" and "intensity" are difficult to define and measure, but should be considered key determinants of providers' performance improvement, however, it should be emphasized that the relationship between intensity of assistance and performance may be non-linear. The PARTner system and the Provider Satisfaction Survey are possible sources of information on the nature of the technical assistance offered by QIOs to providers. Furthermore, measures or scales could be created using detailed descriptions about the methods QIO use to provide technical assistance, the types of information they convey, and the number of times that technical assistance is provided.

III.B Supplementary Short-Term Studies

Our ability to adequately model the IPG selection process and to define and measure key QIO- and provider-specific variables, such as interaction with the QIO, the intensity of technical support and provider “motivation,” limits the ability to conduct a rigorous evaluation of the QIO program. To restructure the program without considering its impact could be costly and, without baseline information on performance, it would be impossible to determine the cost-effectiveness of restructuring. Therefore, we acknowledge the shortcomings of this evaluation option, but believe that many of these limitations could be addressed over time, through investments in short and mid-term studies and additional data collection.

- **Short-term study on IPG selection processes:** There is a dearth of information on the mechanisms that drive inclusion (from the perspective of a QIO) or participation (from the perspective of the provider) in an IPG. A more complete understanding of this relationship is necessary to fully specify the models described above and to accurately control for selection bias in estimating differences in quality improvement for IPG and non-IPG providers. Among the options for better understanding the selection process: (1) interviews could be conducted with QIO staff and providers to understand the criteria that QIOs use to identify IPG candidates and why certain providers opt in or out of the opportunity to participate in an IPG; (2) exploratory secondary data analyses could be conducted to assess how IPG and non-IPG providers differ on basic structural and organizational measures; and (3) the Provider Satisfaction Survey could be modified to collect information on provider-level characteristics that may drive IPG participation, such as motivation and infrastructure availability.
- **Short-term study on types and intensity of QIO interventions:** Scant data exist on the range of technical assistance offered by QIOs and little has been done to characterize the intensity and frequency of QIO interactions with providers. In the short-term, investments in developing measures or scales by which to categorize QIO technical assistance, both in terms of substance and intensity, will further our ability to evaluate the QIO program. Two options for gathering information to develop such a scale include: (1) semi-structured interviews with QIOs and providers to catalog the types of technical assistance strategies and interventions that are employed across all QIOs, and to ascertain whether certain provider or environmental factors influence the decision to use certain types of assistance over others; and (2): the CMS Provider Survey could be modified to gather detailed information on the nature and intensity of specific QIO interventions.

III.C Designs for Evaluating the Special Studies Program

During the 7th SOW, CMS spending on the Special Studies Program amounted to more than \$130 million, of which approximately \$67 million was allocated to QIOSC contracts, which are considered a separate type of special study. Despite the amount dedicated to the Special Studies Program, little is known about how the results of special studies or the assistance provided by QIOSCs support QIO functions or advance the quality of care for Medicare beneficiaries.

- **Special studies:** In the short term, an inventory of key pieces of information on special studies could be developed to support long-term evaluation activities. Through interviews with and a

survey of QIOs, and using CMS administrative data, information could be collected on the status of special studies in the 9th SOW, including special study results, dissemination methods, and target audiences. Building on the information gathered for the inventory, the case study approach may be employed to compare special studies that have been deemed to produce a “good return on investment” to those deemed to produce a “poor return on investment.” Interviews with CMS staff and surveys of QIOs and providers also may provide useful information that speaks to the value that special studies add to the QIO program.

- **QIOSCs:** Similar to the data collection methods used for evaluating special studies, an environmental scan and site visits/interviews with QIOs could be used to gather information on the types and levels of engagement between QIOs and QIOSCs (both topic-specific and cross-cutting). As part of site visits, semi-structured interviews with QIOSC staff could be conducted to gather more detailed information on the nature of the QIO-QIOSC relationship and how QIOSCs attempt to support QIOs. Finally, it may be desirable to invest resources in developing a QIO “engagement scale”, which – by combining information on the substance or nature of technical assistance obtained from QIOSCs with information on the intensity of assistance received – could estimate the level of support QIOSCs provide to specific QIOs. Having developed this scale, collection of data to estimate QIOSC-QIO “scores” could be obtained on an on-going basis by requiring QIOSCs or QIOs to systematically compile and submit data on these interactions to CMS.

III.D Designs for Evaluating Technical Assistance Approaches

Little is known about 1) which approaches for “delivering” technical assistance and 2) the types of content that comprise assistance are most effective in driving quality improvement in particular settings and with particular types of providers. In the short term, semi-structured interviews with QIOs and IPG providers should be conducted to better understand the methods used by QIOs to deliver assistance, the substantive information that is conveyed, and the factors that drive the selection of different methods of assistance. Assuming that issues of confidentiality are addressed, “shadowing” QIO staff as they conduct site visits, seminars, or other training activities could provide an in-depth view that may be unavailable from interviews alone.

CMS’ special study mechanism offers the opportunity to engage QIOs in the study of the effectiveness of technical assistance using more robust, randomized case control, cross-over designs. At minimum, such an approach would examine three models of technical assistance – consultative, collaborative, and provider pay-for-performance – with randomization occurring at either the IPG or QIO level. It should be noted that investments in analyzing alternative approaches are best spent on subtasks for which there is large variation in performance as opposed to those with little variation.

III.E Designs for Extending Support to Poor-Performing and Less Motivated Providers

Project staff and the technical expert panel emphasized the impact that CMS policies governing the QIO program may have on the program’s effectiveness. Of specific interest was the question: *Does the QIO program target the appropriate provider population and, if not, should CMS re-focus requirements to*

encourage QIOs to work with providers who may benefit the most from technical assistance, such as poor performers or providers who lack motivation to engage in quality improvement activities and/or work with QIOs? Through the special study mechanism CMS could empower QIOs to develop alternative approaches for selecting and motivating providers, as well as exploring creative solutions to work with providers to achieve selected performance objectives.

- **Extending support to poor-performing providers:** During the 8th SOW, QIOs were required to offer technical assistance to a maximum of 3 nursing homes that were determined by the State Survey Agency to be “persistently poor” performing homes. In the short term, a case study approach of QIO experiences working with these nursing homes (and, in turn, the nursing homes’ experiences working with QIOs) could be implemented to gather information relevant for evaluating whether CMS should re-focus requirements to encourage QIOs to work with poor performers.
- **Extending support to less motivated providers:** If, as suggested by some members of the TEP, provider motivation is endogenous, it could potentially be influenced by QIOs. As a special study at the outset of the 9th SOW, QIOs could be given the latitude to explore various strategies, including those involving financial and non-financial incentives, to ascertain which ones are most effective in motivating providers to work with QIOs to achieve selected quality improvement objectives. After having identified subsets of providers in selected task areas (e.g., nursing home, home health) randomized case-control studies may be conducted to determine whether selected approaches are more or less effective.

III.F Designs for Evaluating CMS Performance Targets

It is unclear how CMS identifies its quality improvement benchmarks. During site visits, many QIOs reported that they could not meet CMS performance targets because they were “unrealistic” – in large part because there is no known scientific evidence to suggest current targets could be achieved within the time frame used to evaluate performance and, in some cases, because QIOs believed that particular characteristics of their beneficiary or provider population made these targets less feasible or appropriate. Overall, CMS’ approach for setting performance measures and targets must become more transparent if QIOs are to understand more fully the goals they are expected to achieve. To this end:

- Interviews with CMS staff could be conducted to determine the process by which performance targets are set;
- Relevant literature could be reviewed to document ranges of performance improvement that have been achieved by specific types of providers in given time frames;
- In cases where evidence is unavailable to support CMS’ benchmarks, tasks with the greatest variation could be identified for more in-depth investigation, such as through case studies of high- and low-performing QIOs to determine which characteristics are associated with variation in performance; and
- A consensus panel should be convened to review evidence from the literature and from QIO experiences to assist CMS in establishing more realistic performance measures and targets.

IV. Options for Future Evaluation

CMS has made significant investments in the QIO program. Therefore, we recommend that an ongoing or continuous process for evaluating the program would best ensure that funds are spent in the most cost-effective manner. Ideally, the data collection tools and processes used to evaluate a program and the program itself are developed concurrently. Otherwise, the information necessary to adequately conduct the evaluation may not be available at the time the evaluation occurs. Evaluation of the 8th SOW will require the use of retrospective approaches and, therefore, may suffer from the same methodological shortcomings as previous studies. Moving towards the 9th SOW and beyond, prospective, rigorous approaches may be feasible if the data and systems necessary to conduct these evaluations are in place. Therefore, we propose the following three major options:

- (1) **Assess CMS Data Systems & Develop Systems for On-going Evaluation of the QIO Program:** To facilitate future evaluations, a thorough review of CMS's QIO data systems could first be conducted, followed by the development, validation, and incorporation of appropriate data collection tools into the QIO program prior to the start of the SOW – particularly with an eye toward minimizing data lags.
- (2) **Address Limitations in Access to Provider Identifying Data:** In conducting this project, access to data was limited due regulations which prohibit the release of data with provider identifiers; this includes information on whether a provider is a member of an IPG. In an effort to foster and facilitate evaluation of the QIO program, consideration must be given to whether or not such stringent provider confidentiality requirement continues to be needed.
- (3) **Maintain Transparency in Designing and Conducting Evaluation.** The success of an evaluation will, to a great extent, depend on the ability of the evaluator to gain the cooperation of and work effectively with CMS, the QIOs, and providers, all of whom may be asked to contribute information on their operations, collect or submit data, and participate in specific evaluation projects. For these reasons, we highly recommend that the evaluator maintain transparency in designing and conducting the evaluation.